On the inverse of the discrepancy for infinite dimensional infinite sequences

Christoph Aistleitner*

Abstract

In 2001 Heinrich, Novak, Wasilkowski and Woźniakowski proved the upper bound $N^*(s,\varepsilon) \leq c_{\text{abs}} s \varepsilon^{-2}$ for the inverse of the star discrepancy $N^*(s,\varepsilon)$. This is equivalent to the fact that for any $N \geq 1$ and $s \geq 1$ there exists a set of N points in the s-dimensional unit cube whose star-discrepancy is bounded by $c_{\text{abs}}\sqrt{s}/\sqrt{N}$. Dick showed that there exists a double infinite matrix $(x_{n,i})_{n\geq 1, i\geq 1}$ of elements of $[0,1]$ such that for any N and s the star discrepancy of the s-dimensional N-element sequence $((x_{n,i})_{1\leq i\leq s})_{1\leq n\leq N}$ is bounded by

$$\frac{c_{\text{abs}}\sqrt{s \log N}}{\sqrt{N}}.$$

In the present paper we show that this upper bound can be reduced to $c_{\text{abs}}\sqrt{s}/\sqrt{N}$, which is (up to the value of the constant) the same upper bound as the one obtained by Heinrich *et al.* in the case of fixed N and s.

1 Introduction and statement of results

The star discrepancy $D_N^*(x_1,\dots,x_N)$ of a sequence of points $(x_1,\dots,x_N)$ from the s-dimensional unit cube is defined as

$$D_N^*(x_1,\dots,x_N) = \sup_{I\subset[0,1]^s} \left| \lambda(I) - \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_I(x_n) \right|.$$

Here the supremum is taken over all axis-parallel boxes I which are contained in $[0,1]^s$ and have a vertex in the origin, and λ denotes the Lebesgue measure. The so-called *Quasi-Monte Carlo method* is based on the fact that point sequences having small discrepancy can be used for numerical integration. There exist many constructions of point sequences having small discrepancy, such as for example Halton sequences, Sobol sequences, etc. The discrepancy of the first N elements of such sequences (in dimension s) is bounded by $\mathcal{O}\left((\log N)^s N^{-1}\right)$, which is close to the optimal asymptotic order. However, discrepancy bounds of this type are only useful if the number of points N is very large in comparison with the dimension s. For this reason the notion of the *inverse of the discrepancy* was introduced: $N^*(s,\varepsilon)$ denotes the smallest possible number of points in the s-dimensional unit cube which have star discrepancy

*Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: `aistleitner@math.tugraz.at`. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23.

MSC2010: 65C05, 65D32, 11K81

not exceeding ε. By a profound result of Heinrich, Novak, Wasilkowski and Woźniakowski [13] we have

$$N^*(s,\varepsilon) \leq c_{\text{abs}} s \varepsilon^{-2},$$

which is equivalent to the fact that for any N and s there exists a sequence of N points in $[0,1]^s$ whose discrepancy is bounded by $c_{\text{abs}}\sqrt{s}/\sqrt{N}$ (c_{abs} denotes absolute constants, not always the same). Hinrichs [14] proved

$$N^*(s,\varepsilon) \geq c_{\text{abs}} s \varepsilon^{-1},$$

and thus the inverse of the star-discrepancy depends linearly on the dimension s. The dependence on ε is still an open problem.

The proof of Heinrich *et al.* uses a combinatorial result of Haussler, together with a result of Talagrand on empirical processes. In fact, what Heinrich *et al.* actually proved is the following: let $X_1, \dots, X_N$ be a sequence of independent, identically distributed (i.i.d.) $[0,1]^s$-uniformly-distributed random variables. Then with positive probability the discrepancy of $(X_1, \dots, X_N)$ is bounded by

$$c_{\text{abs}}\sqrt{s}/\sqrt{N}. \tag{1}$$

Extending this method, Dick [6] proved the existence of a (double infinite) matrix $(x_{n,i})_{n\geq 1, i\geq 1}$ of numbers $x_{n,i} \in [0,1]$ such that for any $N \geq 1$ and $s \geq 1$ the discrepancy of the s-dimensional N-element sequence $((x_{1,1}, \dots, x_{1,s}), \dots, (x_{N,1}, \dots, x_{N,s}))$ is bounded by

$$c_{\text{abs}}\sqrt{s \log N}/\sqrt{N}. \tag{2}$$

This means that there exist point sequences having small discrepancy, which can be extended both in the dimension s and the number of points N. This can be a significant advantage in applications. More precisely, Dick proved that a randomly generated double infinite matrix satisfies the aforementioned discrepancy bound with positive probability. This asymptotic upper bound contains an additional logarithmic factor in comparison with the estimate (1) for fixed N and s. However, it is clear that a entirely randomly generated matrix cannot achieve the bound (1) uniformly in N and s with positive probability, since by the Chung-Smirnov law of the iterated logarithm (see [20, p. 504]) already for the one-dimensional projections $(x_{1,1}, \dots, x_{N,1})$ of such a matrix we have

$$\limsup_{N\to\infty} \frac{N D_N^*(x_{1,1}, \dots, x_{N,1})}{\sqrt{2N \log\log N}} = \frac{1}{2} \qquad \text{a.s.}$$

(the same asymptotic result holds for all s-dimensional projections for fixed s, see [19, Corollary 4.1.2]). Dick's result has been slightly improved by Doerr, Gnewuch, Kritzer and Pillichshammer [8], who obtained $c_{\text{abs}}\sqrt{s\log(1+N/s)}/\sqrt{N}$ instead of (2). In [3] we have further improved this upper bound to $\sqrt{c_{\text{abs}} s + c_{\text{abs}} \log\log N}/\sqrt{N}$, which is essentially the optimal upper bound which holds for a completely randomly generated matrix with positive probability.

The purpose of the present paper is to prove the existence of a double infinite matrix $(x_{n,i})_{n\geq 1, i\geq 1}$ such that the discrepancy of all its $N \times s$-dimensional projections is bounded by

$$c_{\text{abs}}\sqrt{s}/\sqrt{N}, \qquad \text{for all } s \geq 1 \text{ and } N \geq 1.$$

This is the same upper bound as the one obtain by Heinrich *et al.* in the case of fixed N and s. Since such an upper bound can not be achieved by a entirely randomly generated matrix, we will use a hybrid construction, which consists of both random and deterministic components. More precisely, elements $x_{n,i}$ will be chosen randomly if i is relatively large in comparison with n, while they will be chosen as coordinates of points of an appropriate deterministic low-discrepancy sequence if n is very large in comparison with i.

For a comprehensive treatment of problems and results concerning the inverse of the discrepancy and feasibility of high-dimensional numerical integration by Quasi-Monte Carlo methods we refer the reader to Gnewuch's survey article [12] and to the books of Novak and Woźniakowski [17, 18]. For a general background on discrepancy theory we refer to the monographs of Chazelle [5], Drmota and Tichy [9] and Matoušek [15].

The main result of the present paper is the following Theorem 1.

Theorem 1 *There exists a (double infinite) matrix $(x_{n,i})_{n\geq 1, i\geq 1}$ of elements $x_{n,i} \in [0,1]$, $n \geq 1, i \geq 1$, such that for all $N \geq 1$ and $s \geq 1$ the star-discrepancy D_N^* of the s-dimensional N-element sequence*

$$((x_{1,1},\dots,x_{1,s}),\dots,(x_{N,1},\dots,x_{N,s}))$$

is bounded by

$$D_N^* \leq \frac{234\sqrt{s}}{\sqrt{N}}.$$

2 Preliminaries

Lemma 1 is a simple consequence of [7, Theorem 3.36].

Lemma 1 *For the star-discrepancy of the first N elements of a van der Corput sequence in base $p_1 = 2$ we have the upper bound*

$$D_N^* \leq \frac{1}{\sqrt{N}}, \qquad \text{for } N \geq 1.$$

Lemma 2 ([16, Theorem 3.6]) *Let P denote the first N elements of a Halton sequence in d dimensions, constructed with coprime integers $b_1,\dots,b_d$. Then for all $N \geq 1$,*

$$D_N^*(P) \leq \frac{d}{N} + \frac{1}{N}\prod_{i=1}^{d}\left(\frac{b_i-1}{2\log b_i}\log N + \frac{b_i+1}{2}\right).$$

In the statement of Lemma 2, and throughout the rest of this paper, "log" denotes the natural logarithm.

In our proof will we choose the first d primes $p_1,\dots,p_d$ for the construction of a d-dimensional Halton sequence. In this case we get the following corollary of Lemma 2. We use the fact that for the size of the i-th prime p_i we have $i\log i \leq p_i \leq 1 + 3/2 i \log i$ for $i \geq 2$ (see, for example, [4, Theorem 8.8.4]).

Corollary 1 *Let P denote the first N elements of a Halton sequence in d dimensions, with bases $p_1, \dots, p_d$. Then for any $d \geq 2$*

$$D_N^*(P) \quad \leq \quad \frac{\sqrt{d}}{\sqrt{N}}$$

provided $N \geq 2^{(2^{d+2})}$.

Proof: By Lemma 2 and the subsequent remark we have

$$\begin{aligned} D_N^*(P) \quad &\leq \quad \frac{d}{N} + \frac{\log N}{N} \prod_{i=2}^{d} \left(\frac{3i \log N}{4} + 2i \log i \right) \\ &\leq \quad \frac{d}{N} + \frac{\log N}{N} \prod_{i=2}^{d} (i \log N) \\ &= \quad \frac{d}{N} + \frac{d!(\log N)^d}{N}. \end{aligned}$$

for $N \geq 2^{(2^{d+2})}$. To prove the corollary, it is sufficient to show that

$$\frac{\sqrt{d}}{\sqrt{N}} + \frac{d!(\log N)^d}{\sqrt{dN}} \leq 1 \tag{3}$$

for $N \geq 2^{(2^{d+2})}$. Assume that d is fixed. The derivative of the function $((\log N)^d)/\sqrt{N}$ is zero for $N = e^{2d}$, and negative for $N \geq 2^{(2^{d+2})} \geq e^{2d}$. Thus it is sufficient to show (3) for $N = 2^{(2^{d+2})}$. One can easily check that (3) is true for $N = 2^{(2^{d+2})}$ and $d \in \{2, 3, 4, 5\}$. For $d \geq 6$ we have

$$\begin{aligned} \frac{\sqrt{d}}{\sqrt{N}} + \frac{d!(\log N)^d}{\sqrt{dN}} \quad &\leq \quad \frac{(\log N)^{2d})}{\sqrt{N}} \\ &\leq \quad 2^{2d^2+4d-2^{d+1}} < 1 \end{aligned}$$

for $N = 2^{(2^{d+2})}$, which proves the corollary.

Lemma 3 (Maximal Bernstein inequality; see e.g. [10, Lemma 2.2]) *For a sequence $Z_1, \dots, Z_N$ of i.i.d. random variables having mean zero and variance σ^2, and satisfying $|Z_i| \leq 1$, we have for $t \geq 0$*

$$\mathbb{P} \left(\max_{1 \leq M \leq N} \left| \sum_{n=1}^{M} Z_n \right| > t \right) \leq 2e^{-t^2/(2N\sigma^2 + 2t/3)}$$

Lemma 4 (Triangle inequality for discrepancies; see e.g. [7, Proposition 3.16]) *Let $y_1, \dots, y_N$ be points in $[0,1]^s$. Then for any $1 \leq M < N$*

$$D_N^*(y_1, \dots, y_N) \leq \frac{M D_M^*(y_1, \dots, y_M)}{N} + \frac{(N-M) D_{N-M}^*(y_{M+1}, \dots, y_N)}{N}$$

and

$$D_{N-M}^*(y_{M+1}, \dots, y_N) \leq \frac{N D_N^*(y_1, \dots, y_N)}{N-M} + \frac{M D_M^*(y_1, \dots, y_M)}{N-M}.$$

Throughout this paper, for points $v, w \in [0,1]^s$ we will write $v \leq w$ if this inequality holds coordinatewise. Furthermore, we will write 0 for the s-dimensional vector $(0, \ldots, 0)$, and $[0, w]$ for the set $\{v \in [0,1]^s : 0 \leq v \leq w\}$.
For some number $\delta > 0$ a set Γ of points in $[0,1]^s$ is called a *δ-cover* if for every $x \in [0,1]^s$ there exist points $v, w \in \Gamma \cup \{0\}$ such that $v \leq x \leq w$ and $\lambda([0,w]) - \lambda([0,v]) \leq \delta$. Similarly, a set Δ of elements of $[0,1]^s \times [0,1]^s$ is called a *δ-bracketing cover* if for every pair $(v, w) \in \Delta$ the estimate $\lambda([0,w]) - \lambda([0,v]) \leq \delta$ holds, and if for every $x \in [0,1]^s$ there exists $(v, w) \in \Delta$ such that $v \leq x \leq w$. These two notions are closely related, and they both are very useful for reducing the calculation of the star discrepancy from evaluating a supremum over all possible intervals to evaluating a maximum over a finite set of intervals. For details on the definitions and properties of δ-covers and δ-bracketing covers, see [11, 12].

Lemma 5 ([11, Theorem 1.15]) *For any $s \geq 1$ and $\delta > 0$, there exist a δ-cover Γ and a δ-bracketing cover Δ of cardinality at most $(2e)^s \left(\delta^{-1} + 1\right)^s$, respectively.*

3 Proof of Theorem 1

For a number $b \geq 2$, and any $n \geq 1$, let

$$\nu_0 + \nu_1 b + \nu_2 b^2 \ldots$$

be the (finite) b-adic expansion of n, and set

$$\varphi_b(n) = \frac{\nu_0}{b} + \frac{\nu_1}{b^2} + \frac{\nu_2}{b^3} + \ldots$$

The function $\varphi_b : \mathbb{N} \to [0,1)$ is called the *(b-adic) radical inverse function.* For $n \geq 1,\ i \geq 1$, set

$$q_{n,i} = \varphi_{p_i}(n),$$

where p_i is the i-th prime. Then the points $(q_{n,1}, \ldots, q_{n,d})_{1 \leq n \leq N}$ are the first N elements of the d-dimensional Halton sequence with bases $p_1, \ldots, p_d$. For such point sets we have the discrepancy estimates in Lemma 1 and Corollary 1 below.

Let $(X_{n,i})_{n \geq 1, i \geq 1}$ be an array of i.i.d. random variables, all of which have uniform distribution on $[0,1]$. For $n \geq 1, i \geq 1$, set

$$x_{n,i} = \begin{cases} q_{n,i} & \text{if } i = 1 \text{ or } 2^{(2^{i+2})} < n \\ X_{n,i} & \text{if } i \geq 2 \text{ and } 2^{(2^{i+2})} \geq n. \end{cases} \tag{4}$$

This means that the matrix $(x_{n,i})_{n \geq 1, i \geq 1}$ has both random and deterministic components, depending on the relation of the indices n and i. We will write $D_N^s(x_{n,i})$ for the star-discrepancy of the N-element set of s-dimensional points

$$\Big\{(x_{1,1}, \ldots, x_{1,s}), \ldots, (x_{N,1}, \ldots, x_{N,s})\Big\},$$

and, for $0 \leq M < N$, we will write $D_{M,N}^s(x_{n,i})$ for the star-discrepancy of the set of $N - M$ points

$$\Big\{(x_{M+1,1}, \ldots, x_{M+1,s}), \ldots, (x_{N,1}, \ldots, x_{N,s})\Big\}.$$

For $m \geq 1$ and $s \geq 1$, set

$$A_{m,s} = \left\{ \max_{2^m < M \leq 2^{m+1}} M D_M^s(x_{n,i}) \geq c_{m,s} \sqrt{s} \sqrt{2^{m+1}} \right\},$$

where

$$c_{m,s} = \begin{cases} 163 & \text{if } 2^{s+2} > m \\ 165 & \text{if } 2^{s+2} \leq m. \end{cases}$$

We will show that

$$\mathbb{P} \left(\bigcup_{s=1}^{\infty} \bigcup_{m=1}^{\infty} A_{m,s} \right) < 1. \tag{5}$$

Since on the complement of $(\bigcup_{s=1}^{\infty} \bigcup_{m=1}^{\infty} A_{m,s})$ we have

$$N D_N^s(x_{n,i}) \leq 165 \sqrt{s} \sqrt{2N} \qquad \text{for all} \quad N, s,$$

this clearly proves the existence of a matrix $(z_{n,i})_{n \geq 1, i \geq 1}$ for which

$$D_N^s(z_{n,i}) \leq \sqrt{2} \cdot 165 \frac{\sqrt{s}}{\sqrt{N}} \leq 234 \frac{\sqrt{s}}{\sqrt{N}} \qquad \text{for all } N \geq 1, \ s \geq 1.$$

Thus for the proof of Theorem 1 it remains to show (5).

For $s = 1$ and $m \geq 1$ we have

$$\mathbb{P}(A_{m,s}) = 0 \tag{6}$$

by Lemma 1. Thus we will henceforth always assume that $s > 1$.

Let $s \geq 2$ and $m \geq 1$ be given. Assume that $m \geq 2^{s+2}$. Set $\mu = 2^{s+2}$. Then, by the first part of Lemma 4, for any integer $M \in (2^m, 2^{m+1}]$

$$M D_M^s(x_{n,i}) \leq 2^{\mu} D_{2^{\mu}}^s(x_{n,i}) + (M - 2^{\mu}) D_{2^{\mu},M}(x_{n,i}). \tag{7}$$

The $M - 2^{\mu}$ points

$$\left\{ (x_{2^{\mu}+1,1}, \ldots, x_{2^{\mu}+1,s}), \ldots, (x_{M,1}, \ldots, x_{M,s}) \right\}$$

are purely deterministic, namely the points with index $2^{\mu} + 1, \ldots, M$ of the s-dimensional Halton sequence with bases $p_1, \ldots, p_s$. Thus by Corollary 1 and the second part of Lemma 4

$$\begin{aligned} D_{2^{\mu},M}^s(x_{n,i}) = D_{2^{\mu},M}^s(q_{n,i}) &\leq \frac{2^{\mu} D_{2^{\mu}}^s(q_{n,i}) + M D_M^s(q_{n,i})}{M - 2^{\mu}} \\ &\leq \frac{2^{\mu} \sqrt{s} / \sqrt{2^{\mu}} + M \sqrt{s} / \sqrt{M}}{M - 2^{\mu}} \\ &\leq \frac{\sqrt{s} \sqrt{2^m} + \sqrt{s} \sqrt{2^{m+1}}}{M - 2^{\mu}} \\ &< \frac{2 \sqrt{s} \sqrt{2^{m+1}}}{M - 2^{\mu}}. \end{aligned} \tag{8}$$

By definition we have

$$\left\{ 2^{\mu} D_{2^{\mu}}^s(x_{n,i}) \geq 163 \sqrt{s} \sqrt{2^{\mu}} \right\} \subset A_{\mu-1,s}.$$

Thus by (7) and (8)

$$\begin{aligned} A_{m,s} \backslash A_{\mu-1,s} &\subset \left\{ \max_{2^m < M \le 2^{m+1}} M D_M^s(x_{n,i}) \ge 165 \sqrt{s} \sqrt{2^{m+1}} \right\} \backslash \left\{ 2^\mu D_{2^\mu}^s(x_{n,i}) \ge 163 \sqrt{s} \sqrt{2^\mu} \right\} \\ &\subset \left\{ \max_{2^m < M \le 2^{m+1}} (M - 2^\mu) D_{2^\mu,M}^s(x_{n,i}) > 2\sqrt{s}\sqrt{2^{m+1}} \right\} = \emptyset, \end{aligned}$$

and consequently we have for all $m \ge \mu$

$$\mathbb{P}(A_{m,s} \backslash A_{\mu-1,s}) = 0.$$

Together with (6) this implies

$$\mathbb{P}\left(\bigcup_{s=1}^{\infty} \bigcup_{m=1}^{\infty} A_{m,s} \right) = \mathbb{P}\left(\bigcup_{s=2}^{\infty} \bigcup_{1 \le m < 2^{s+2}} A_{m,s} \right), \tag{9}$$

and to prove (5) it remains to estimate the probabilities of the sets $A_{m,s}$ for $s \ge 2$ and $1 \le m < 2^{s+2}$.

Assume that $s \ge 2$ and $m < 2^{s+2}$. Additionally we can assume that

$$\frac{\sqrt{s}}{\sqrt{2^{m+1}}} \le \frac{1}{64}, \tag{10}$$

since otherwise trivially $A_{m,s} = \emptyset$. Set $\hat{k}(m) = \max\{k \ge 1 : \; 2^{k+2} \le m\}$. If $m \ge 16$, we set $\mathrm{L} = 2^{2^{\hat{k}(m)+2}}$. If $m < 16$, we set $L = 0$. Note that the value of L depends on m, and that $L \le 2^m$. We define sets

$$\begin{aligned} G(m,s) &= \begin{cases} \left\{ L D_L^s(x_{n,i}) \ge 82 \sqrt{s} \sqrt{L} \right\} & \text{if } L > 0 \\ \emptyset & \text{if } L = 0 \end{cases} \\ H(m,s) &= \left\{ \max_{L < M \le 2^{m+1}} (M - L) D_{L,M}^s(x_{n,i}) \ge 81 \sqrt{s} \sqrt{2^{m+1}} \right\}. \end{aligned}$$

Then we claim

$$A_{m,s} \subset G(m,s) \cup H(m,s).$$

This is clear for $L = 0$. For $L > 0$ (which implies $m \ge 16$ and $\hat{k}(m) \ge 3$) we have by the first part of Lemma 4,

$$\max_{2^m < M \le 2^{m+1}} M D_M^s(x_{n,i}) \le \max_{2^m < M \le 2^{m+1}} \left(L D_L^s(x_{n,i}) + (M - L) D_{L,M}^s(x_{n,i}) \right)$$

and thus on $(G(m,s) \cup H(m,s))^C$

$$\begin{aligned} \max_{2^m < M \le 2^{m+1}} M D_M^s(x_{n,i}) &\le 82 \sqrt{s} \sqrt{L} + 81 \sqrt{s} \sqrt{2^{m+1}} \\ &\le 163 \sqrt{s} \sqrt{2^{m+1}}. \end{aligned}$$

Note that by definition for any $k \ge 3$ all the sets $G(m,s)$ for $m = 2^{k+2}, \dots, 2^{k+3} - 1$ are equal. Thus it is sufficient to consider the sets $G(m,s)$ for m of the form

$$m = 2^{k+2}, \qquad \text{for } k \ge 3. \tag{11}$$

In the case $m = 2^5 = 16$ we have $L = 2^{16}$, and

$$G(m,s) \subset \left\{ 2^{16} D^s_{2^{16}}(x_{n,i}) \geq 81\sqrt{s}\sqrt{L} \right\} \subset H(15,s).$$

If (11) holds for some $k \geq 4$, then we have $\hat{k}(m) = k$ and $L = 2^{(2^{k+2})}$. Note that in this case $\hat{k}(m-1) = k-1$ and $2^{(2^{\hat{k}(m-1)})} = 2^{(2^{k+1})} = \sqrt{L}$. We have

$$LD^s_L(x_{n,i}) \quad \leq \quad 2^{(2^{k+1})} + \left(2^{(2^{k+2})} - 2^{(2^{k+1})} \right) D^s_{2^{(2^{k+1})}, 2^{(2^{k+2})}}(x_{n,i}),$$

due to the first part of Lemma 4, which implies

$$\begin{aligned} G(m,s) \quad &\subset \quad \left\{ \left(2^{(2^{k+2})} - 2^{(2^{k+1})} \right) D^s_{2^{(2^{k+1})}, 2^{(2^{k+2})}}(x_{n,i}) \geq 82\sqrt{s}\sqrt{L} - \sqrt{L} \right\} \\ &\subset \quad \left\{ \left(2^{(2^{k+2})} - 2^{(2^{k+1})} \right) D^s_{2^{(2^{k+1})}, 2^{(2^{k+2})}}(x_{n,i}) \geq 81\sqrt{s}\sqrt{L} \right\} \\ &\subset \quad H(m-1,s). \end{aligned}$$

Thus for any s

$$\bigcup_{1 \leq m < 2^{s+2}} A_{m,s} \subset \bigcup_{1 \leq m < 2^{s+2}} (G(m,s) \cup H(m,s)) \subset \bigcup_{1 \leq m < 2^{s+2}} H(m,s), \tag{12}$$

and to estimate the probability of $\bigcup_{1 \leq m < 2^{s+2}} A_{m,s}$ it is sufficient to estimate the probabilities of $H(m,s),\ 1 \leq m < 2^{s+2}$.

Next we will estimate the probability of the sets $H(m,s)$, for fixed $s \geq 2$ and $1 \leq m < 2^{s+2}$. We will use a method which is somewhat similar to that in [2], but in the present case the situation is slightly more complicated. Set

$$K = \lceil (m+1)/2 - (\log_2 s)/2 - 2 \rceil. \tag{13}$$

Then $K \geq 4$ due to (10), and consequently

$$\sqrt{s}\sqrt{2^{m+1}} \leq 2^{m-k} \qquad \text{for any } \ k,\ 1 \leq k \leq K. \tag{14}$$

For $1 \leq k \leq K-1$, let Γ_k denote a 2^{-k}-cover of $[0,1]^s$, for which

$$\#\Gamma_k \leq (2e)^s (2^k + 1)^s \leq \left\{ \begin{array}{ll} (6e)^{ks} & \text{for } k = 1 \\ (2e)^s (\sqrt{5})^{ks} & \text{for } k > 1 \end{array} \right. \tag{15}$$

and let Δ_K denote a 2^{-K}-bracketing cover of $[0,1]^s$ for which

$$\#\Delta_K \leq (2e)^s (2^K + 1)^s \leq (2e)^s (\sqrt{5})^{Ks}. \tag{16}$$

Such covers exist by Lemma 5. For notational convenience we set

$$\begin{aligned} \Gamma_K &= \{ v \in [0,1]^s :\ (v,w) \in \Delta_K \quad \text{for some } w \}, \\ \Gamma_{K+1} &= \{ w \in [0,1]^s :\ (v,w) \in \Delta_K \quad \text{for some } v \}, \end{aligned}$$

and for points $x, y \in [0,1]^s$

$$\overline{[x,y]} := \begin{cases} [0,y] \backslash [0,x] & \text{if } x \neq 0 \\ [0,y] & \text{if } x = 0 \\ \emptyset & \text{if } x = y = 0. \end{cases}$$

Then for an arbitrary point $x \in [0,1]^s$ there exist sets $I_k(x),\ 0 \leq k \leq K$ such that

$$\bigcup_{k=0}^{K-1} I_k(x) \subset [0,x] \subset \bigcup_{k=0}^{K} I_k(x), \tag{17}$$

and each I_k is of the form $\overline{[p_k(x), p_{k+1}(x)]}, 0 \leq k \leq K$, where $p_0 = 0$ and $p_i(x) \in \Gamma_i, 1 \leq i \leq K+1$ (see [1] for details). Furthermore, each set I_k has volume at most 2^{-k}, and as x runs through the whole unit cube $[0,1]^s$ we obtain at most $\#\Gamma_{k+1}$ different sets $I_k(x)$, for $0 \leq k \leq K$. We write S_k for the class of all sets of the form $I_k(x)$ for some $x \in [0,1]^s$, for $0 \leq k \leq K$.

For $m \geq 16$, set $d = \hat{k}(m)$. For $m < 16$ set $d = 1$. Then for any $n \in (L, 2^{m+1}]$ we have from (4) that $x_{n,i} = q_{n,i}$ if $i \leq d$, and $x_{n,i} = X_{n,i}$ if $i > d$. In other words, for all the s-dimensional points in the sequence

$$\big((x_{L+1,1}, \ldots, x_{L+1,s}), \ldots, (x_{2^{m+1},1}, \ldots, x_{2^{m+1},s})\big)$$

the first d coordinates are deterministic and the remaining $s-d$ coordinates are random (note that $s \geq 2$ and $m < 2^{s+2}$ implies $s > d$). This is obvious in the case $m < 16$, when $L = 0$ and only the first coordinate of the points is deterministic. If $m \geq 16$, then $2^{\hat{k}(m)+2} \leq m < 2^{\hat{k}(m)+3}$. Thus for the numbers $n \in \{L+1, 2^{m+1}\}$ we have

$$2^{(2^{\hat{k}(m)+2})} < n \leq 2^{(2^{\hat{k}(m)+3})},$$

which by (4) means that exactly the first $\hat{k}(m)$ coordinates of $x_{n,i}$ are deterministic and the remaining coordinates are random.

For any $k \in \{1, \ldots, K+1\}$ the numbers p_k can be written in the form (u_k, v_k), where $u_k \in [0,1]^d$ and $v_k \in [0,1]^{s-d}$. We define $U_k = [0, u_k],\ V_k = [0, v_k]$. Then $U_k \times V_k = [0, p_k]$. For sets $I_k \in S_k$ and $I_{k-1} \in S_{k-1}$ we write $I_{k-1} \prec I_k$ if there exists an $x \in [0,1]^s$ such that $I_{k-1} = I_{k-1}(x)$ and $I_k = I_k(x)$. For every $I_k \in S_k$ there exists exactly one set $I_{k-1} \in S_{k-1}$ such that $I_{k-1} \prec I_k$, for $1 \leq k \leq K$. Every fixed set I_k uniquely determines sets $I_0 \prec \cdots \prec I_{k-1}$ as well as corresponding values for p_l, u_l, v_l, U_l, V_l for $0 \leq l \leq k$. Moreover, every set $I_k \in S_k,\ 1 \leq k \leq K$ is of the form

$$(U_{k+1} \times V_{k+1}) \backslash (U_k \times V_k) = \left((U_{k+1} \backslash U_k) \times V_{k+1}\right) \cup \left(U_k \times (V_{k+1} \backslash V_k)\right),$$

and every set $I_0 \in S_0$ is of the form $U_1 \times V_1$. Hence

$$\lambda(U_{k+1} \backslash U_k) \cdot \lambda(V_{k+1}) \leq \lambda(I_k) \leq 2^{-k}. \tag{18}$$

A similar decomposition is described in more detail in [2].

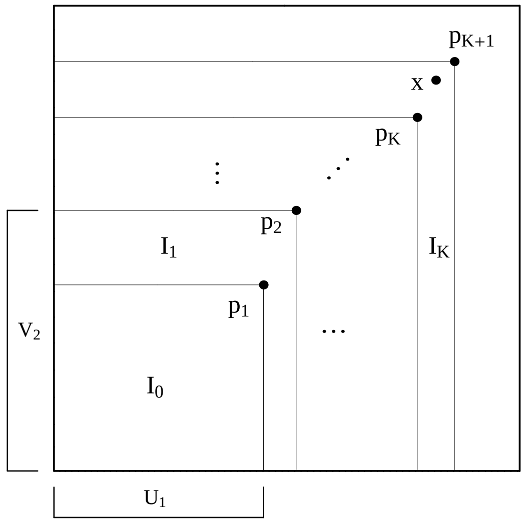

Figure 1: An illustration of the decomposition in the case $d = 1$, $s = 2$. A point $x \in [0,1]^2$ is given and determines points $p_0, p_1, \ldots, p_{K+1}$ and sets $I_0 \prec I_1 \prec \cdots \prec I_K$. Every set I_k, $1 \le k \le K$, is of the form $(U_{k+1} \times V_{k+1}) \backslash (U_k \times V_k) = ((U_{k+1} \backslash U_k) \times V_{k+1}) \cup (U_k \times (V_{k+1} \backslash V_k))$, the set I_0 is of the form $U_1 \times V_1$.

For abbreviation we write

$$x_n = (x_{n,1}, \ldots, x_{n,s}), \quad q_n = (q_{n,1}, \ldots, q_{n,d}), \quad X_n = (X_{n,d+1}, \ldots, X_{n,s}).$$

Then for arbitrary $x \in [0,1]^s$ and $M \in \{L+1, \ldots, 2^{m+1}\}$, by (17),

$$\begin{aligned} & \sum_{n=L+1}^{M} \mathbb{1}_{[0,x]}(x_n) \\ \ge \; & \sum_{n=L+1}^{M} \mathbb{1}_{[0,p_K]}(x_n) \\ = \; & \sum_{n=L+1}^{M} \mathbb{1}_{U_K}(q_n) \cdot \mathbb{1}_{V_K}(X_n) \\ = \; & \sum_{n=L+1}^{M} \mathbb{1}_{U_1}(q_n) \cdot \mathbb{1}_{V_1}(X_n) \\ & + \sum_{k=1}^{K-1} \sum_{n=L+1}^{M} \left(\mathbb{1}_{U_{k+1} \backslash U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}}(X_n) + \mathbb{1}_{U_k}(q_n) \cdot \mathbb{1}_{V_{k+1} \backslash V_k}(X_n) \right), \end{aligned} \tag{19}$$

and similarly

$$\begin{aligned}
&\sum_{n=L+1}^{M} \mathbb{1}_{[0,x]}(x_n) \\
\le\ & \sum_{n=L+1}^{M} \mathbb{1}_{U_1}(q_n) \cdot \mathbb{1}_{V_1}(X_n) \\
&+ \sum_{k=1}^{K} \sum_{n=L+1}^{M} \left(\mathbb{1}_{U_{k+1}\backslash U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}}(X_n) + \mathbb{1}_{U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}\backslash V_k}(X_n) \right). \qquad (20)
\end{aligned}$$

Note that for arbitrary $M \in \{2^m+1, \dots, 2^{m+1}\}$ by Corollary 1 we have

$$(M-L)D^*_{L,M}(q_n) \le LD^*_L(q_n) + MD^*_M(q_n) \le \sqrt{d}\sqrt{L} + \sqrt{d}\sqrt{M} \le 2\sqrt{d}\sqrt{M}. \qquad (21)$$

Additionally Corollary 1 implies for any $k \in \{1, \dots, K\}$ that

$$\sum_{n=L+1}^{2^{m+1}} \mathbb{1}_{U_{k+1}\backslash U_k}(q_n) \le \sum_{n=1}^{2^{m+1}} \mathbb{1}_{U_{k+1}\backslash U_k}(q_n) \le 2^{m+1}\lambda\left(U_{k+1}\backslash U_k\right) + 2\sqrt{d}\sqrt{2^{m+1}}, \qquad (22)$$

and similarly

$$\sum_{n=L+1}^{2^{m+1}} \mathbb{1}_{U_k}(q_n) \le 2^{m+1}\lambda\left(U_k\right) + 2\sqrt{d}\sqrt{2^{m+1}}. \qquad (23)$$

Thus by Lemma 3 as well as (14), (18) and (22) for every $t > 0$ and any $k \in \{1, \dots, K\}$,

$$\begin{aligned}
&\mathbb{P}\left(\max_{L+1 \le M \le 2^{m+1}} \left| \sum_{n=L+1}^{M} \left(\mathbb{1}_{U_{k+1}\backslash U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}}(X_n) - \mathbb{1}_{U_{k+1}\backslash U_k}(q_n) \cdot \lambda(V_{k+1}) \right) \right| > t \right) \\
=\ & \mathbb{P}\left(\max_{L+1 \le M \le 2^{m+1}} \left| \sum_{\substack{L+1 \le n \le M, \\ q_n \in U_{k+1}\backslash U_k}} \left(\mathbb{1}_{V_{k+1}}(X_n) - \lambda(V_{k+1}) \right) \right| > t \right) \\
\le\ & 2\exp\left(- \frac{t^2}{2\left(\sum_{\substack{L+1 \le n \le 2^{m+1}, \\ q_n \in U_{k+1}\backslash U_k}} 1 \right) \left(\lambda(V_{k+1})\left(1 - \lambda(V_{k+1})\right) \right) + 2t/3} \right) \\
\le\ & 2\exp\left(- \frac{t^2}{2\left(2^{m+1}\lambda\left(U_{k+1}\backslash U_k\right)\right) + 2\sqrt{d}\sqrt{2^{m+1}})\lambda(V_{k+1}) + 2t/3} \right) \\
\le\ & 2\exp\left(- \frac{t^2}{2^{m-k+3} + 2t/3} \right) \qquad (24)
\end{aligned}$$

(here and in the sequel we write $\exp(x)$ for e^x). Similarly, we obtain using (23) instead of (22)

$$\begin{aligned}
&\mathbb{P}\left(\max_{L+1 \le M \le 2^{m+1}} \left| \sum_{n=L+1}^{M} \left(\mathbb{1}_{U_k}(q_n) \cdot \mathbb{1}_{V_{k+1}\backslash V_k}(X_n) - \mathbb{1}_{U_k}(q_n) \cdot \lambda(V_{k+1}\backslash V_k) \right) \right| > t \right) \\
\le\ & 2\exp\left(- \frac{t^2}{2^{m-k+3} + 2t/3} \right),
\end{aligned}$$

and

$$\mathbb{P}\left(\max_{L+1\le M\le 2^{m+1}}\left|\sum_{n=L+1}^{M}\left(\mathbb{1}_{U_1}(q_n)\cdot\mathbb{1}_{V_1}(X_n)-\lambda(V_1)\mathbb{1}_{U_1}(q_n)\right)\right|>t\right)$$

$$\le\ 2\exp\left(-\frac{t^2}{2^{m+3}+2t/3}\right). \qquad (25)$$

We observe that for $t=8\sqrt{s}\sqrt{k}2^{-k/2}\sqrt{2^{m+1}}$ we have by (14) that

$$2t/3\le\frac{16}{3}2^{m-k}\sqrt{k}2^{-k/2}\le 2^{m-k+3} \qquad (26)$$

and

$$2\exp\left(-\frac{t^2}{2^{m-k+3}+2t/3}\right)\le 2\exp\left(-\frac{t^2}{2^{m-k+4}}\right) \qquad (27)$$

Consequently, due to (24) and (27), we have for $1\le k\le K$

$$\mathbb{P}\left(\max_{L+1\le M\le 2^{m+1}}\left|\sum_{n=L+1}^{M}\left(\mathbb{1}_{U_{k+1}\backslash U_k}(q_n)\cdot\mathbb{1}_{V_{k+1}}(X_n)\ -\ \mathbb{1}_{U_{k+1}\backslash U_k}(q_n)\cdot\lambda(V_{k+1})\right)\right|>8\sqrt{s}\sqrt{k}2^{-k/2}\sqrt{2^{m+1}}\right)$$

$$\le\ 2\exp\left(-\frac{\left(8\sqrt{s}\sqrt{k}2^{-k/2}\sqrt{2^{m+1}}\right)^2}{2^{m-k+4}}\right)$$

$$\le\ 2e^{-8ks}. \qquad (28)$$

Analogously, because of (25) and (27), we conclude

$$\mathbb{P}\left(\max_{L+1\le M\le 2^{m+1}}\left|\sum_{n=L+1}^{M}\left(\mathbb{1}_{U_k}(q_n)\cdot\mathbb{1}_{V_{k+1}\backslash V_k}(X_n)\ -\ \mathbb{1}_{U_k}(q_n)\cdot\lambda(V_{k+1}\backslash V_k)\right)\right|>8\sqrt{s}\sqrt{k}2^{-k/2}\sqrt{2^{m+1}}\right)$$

$$\le\ 2e^{-8ks}, \qquad (29)$$

and

$$\mathbb{P}\left(\max_{L+1\le M\le 2^{m+1}}\left|\sum_{n=L+1}^{M}\left(\mathbb{1}_{U_1}(q_n)\cdot\mathbb{1}_{V_1}(X_n)-\lambda(V_1)\mathbb{1}_{U_1}(q_n)\right)\right|>8\sqrt{s}\sqrt{2^{m+1}}\right)$$

$$\le\ 2e^{-8s}, \qquad (30)$$

where we used (26) with $t=8\sqrt{s}\sqrt{2^{m+1}}$ and the fact that $-t^2/(2^{m+3}+2t/3)\le -8s$ due to (10).

By (15) and (16) the number of exceptional sets in (28) and (29) is bounded by $(2e)^s(\sqrt{5})^{(k+1)s}$, respectively, (as x runs through all possible values in $[0,1]^s$), and the number of exceptional sets in (30) is bounded by $(6e)^s$. Thus by (13), (20), (21), (28), (29) and (30) we have for any $M \in \{2^m+1,\dots,2^{m+1}\}$

$$\begin{aligned}
&\sum_{n=L+1}^{M} \mathbb{1}_{[0,x]}(x_n) \\
\leq\ & \sum_{n=L+1}^{M} \mathbb{1}_{U_1}(q_n)\cdot \mathbb{1}_{V_1}(X_n) \\
&\quad + \sum_{k=1}^{K}\sum_{n=L+1}^{M} \left(\mathbb{1}_{U_{k+1}\setminus U_k}(q_n)\cdot \mathbb{1}_{V_{k+1}}(X_n) + \mathbb{1}_{U_k}(q_n)\cdot \mathbb{1}_{V_{k+1}\setminus V_k}(X_n)\right) \\
\leq\ & \left(\sum_{n=L+1}^{M} \mathbb{1}_{U_{K+1}}(q_n)\lambda(V_{K+1})\right) + 8\sqrt{s}\sqrt{2^{m+1}} + 2\sum_{k=1}^{K} 8\sqrt{s}\sqrt{k}2^{-k/2}\sqrt{2^{m+1}} \\
\leq\ & (M-L)\lambda(I_{K+1}) + 2\sqrt{s}\sqrt{2^{m+1}} + 75\ \sqrt{s}\sqrt{2^{m+1}} \\
\leq\ & (M-L)\left(\lambda([0,x]) + 2^{-K}\right) + 77\sqrt{s}\sqrt{2^{m+1}} \\
\leq\ & (M-L)\lambda([0,x]) + 81\sqrt{s}\sqrt{2^{m+1}}
\end{aligned}$$

and, similarly, using (19) instead of (20),

$$\begin{aligned}
&\sum_{n=L+1}^{M} \mathbb{1}_{[0,x]}(x_n) \\
\geq\ & \sum_{n=L+1}^{M} \mathbb{1}_{U_1}(q_n)\cdot \mathbb{1}_{V_1}(X_n) \\
&\quad + \sum_{k=1}^{K-1}\sum_{n=L+1}^{M} \left(\mathbb{1}_{U_{k+1}\setminus U_k}(q_n)\cdot \mathbb{1}_{V_{k+1}}(X_n) + \mathbb{1}_{U_k}(q_n)\cdot \mathbb{1}_{V_{k+1}\setminus V_k}(X_n)\right) \\
\geq\ & (M-L)\lambda([0,x]) - 81\sqrt{s}\sqrt{2^{m+1}}.
\end{aligned}$$

on a set of probability at least

$$\begin{aligned}
1-(6e)^s\left(2e^{-8s}\right) - 4(2e)^s\sum_{k=1}^{K}(\sqrt{5})^{(k+1)s}e^{-8ks} &\geq 1-2(6e)^s e^{-8s} - 5(10e)^s e^{-8s} \\
&\geq 1-7(10e)^s e^{-8s} \geq 1-2^{-2s-2}
\end{aligned}$$

for all $x\in[0,1]^s$ and $M\in\{2^m+1,\dots,2^{m+1}\}$ (remember that we have assumed $s\geq 2$). In other words, we have shown that

$$\mathbb{P}\left(H(m,s)\right) \leq 2^{-2s-2}.$$

Thus by (12)

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{s\geq 2}\ \bigcup_{1\leq m<2^{s+2}} A_{m,s}\right) &\leq \sum_{s=2}^{\infty}\sum_{m=1}^{2^{s+2}-1}\mathbb{P}(H(m,s))\\
&\leq \sum_{s=2}^{\infty} 2^{-s}\\
&\leq \frac{1}{2}.
\end{aligned}$$

Together with (5) and (9) this proves Theorem 1.

References

[1] C. Aistleitner. Covering numbers, dyadic chaining and discrepancy. *J. Complexity*, 27(6):531–540, 2011.

[2] C. Aistleitner and M. Hofer. Probabilistic error bounds for the discrepancy of mixed sequences. *Monte Carlo Methods Appl.*, to appear.

[3] C. Aistleitner and M. Weimar. Probabilistic star discrepancy bounds for double infinite random matrices. Submitted.

[4] E. Bach and J. Shallit. *Algorithmic number theory. Vol. 1.* Foundations of Computing Series. MIT Press, Cambridge, MA, 1996.

[5] B. Chazelle. *The discrepancy method.* Cambridge University Press, Cambridge, 2000.

[6] J. Dick. A note on the existence of sequences with small star discrepancy. *J. Complexity*, 23(4-6):649–652, 2007.

[7] J. Dick and F. Pillichshammer. *Digital nets and sequences. Discrepancy theory and quasi-Monte Carlo integration.* Cambridge: Cambridge University Press. xvii, 600 p., 2010.

[8] B. Doerr, M. Gnewuch, P. Kritzer, and F. Pillichshammer. Component-by-component construction of low-discrepancy point sets of small size. *Monte Carlo Methods Appl.*, 14(2):129–149, 2008.

[9] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1997.

[10] U. Einmahl and D. M. Mason. Some universal results on the behavior of increments of partial sums. *Ann. Probab.*, 24(3):1388–1407, 1996.

[11] M. Gnewuch. Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy. *J. Complexity*, 24(2):154–172, 2008.

[12] M. Gnewuch. Entropy, randomization, derandomization, and discrepancy. In *Monte Carlo and quasi-Monte Carlo methods 2010.* Springer, to appear.

[13] S. Heinrich, E. Novak, G. W. Wasilkowski, and H. Woźniakowski. The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.*, 96(3):279–302, 2001.

[14] A. Hinrichs. Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. *J. Complexity*, 20(4):477–483, 2004.

[15] J. Matoušek. *Geometric discrepancy*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2010.

[16] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[17] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume I: Linear information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.

[18] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.

[19] W. Philipp. *Mixing sequences of random variables and probablistic number theory*. American Mathematical Society, Providence, R. I., 1971. Memoirs of the American Mathematical Society, No. 114.

[20] G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.